

### REMARKS

In response to the Office Action dated October 25, 2010 (hereinafter "the Office Action"), the Assignee (Nuance Communications, Inc.) respectfully requests reconsideration. Claims 1-15 and 17-27 were previously pending in this application. By this response, no claims are amended, cancelled, nor added. As a result, claims 1-15 and 17-27 are pending for examination with claims 1, 6, 11, 17, 21, and 24 being independent.

#### Objection to the Specification

The Office Action (p. 4) objects to the specification as purportedly failing to provide proper antecedent basis for the claimed subject matter. The Office Action states, "There is also no suggestion of a computer readable medium such as RAM, ROM, etc." and "There is also no suggestion of a computer program product and computer memory such as program code, software stored on hardware, ROM, RAM, hard drive, etc." Assignee respectfully traverses the objection.

FIG. 1 of the specification illustrates computing apparatus with processor 108. The specification describes handling of various types of data and model creation carried out with processor 108 (see paragraphs [0015] – [0019] in reference to the published application No. 2005/0049872). Paragraph [0019] describes **storing** and searching of acoustic models. Paragraph [0020] notes, "It should be remarked that the logical operations shown may be implemented (1) as a **sequence of computer executed steps running on a computing system** and/or (2) as interconnected machine modules within the computing system." At least these sections of the specification describe, implicitly or explicitly, and/or suggest computer readable medium, computer program product, and computer memory such as program code. At least in view of these sections, "the meaning of the terms in the claims may be ascertainable by reference to the description" as required by 37 CFR 1.75(d)(1) and MPEP 608.01(o) and cited in the Office Action. Additionally, the claim terms "computer readable medium," "computer program product," and "computer memory" would have been understood by one of ordinary skill in the art at the time of the invention, since such terms were in common use at that time. For at least these reasons, withdrawal of the objection to the specification is respectfully requested.

Claim Rejections Under 35 U.S.C. § 112, First Paragraph

The Office Action (p. 5) rejects claims 1-5, 11-15, 17-20, and 24-27 as purportedly failing to comply with the written description requirement. Assignee respectfully traverses the rejection.

For reasons that should be clear from at least FIG. 1, and paragraphs [0015] – [0020] discussed above in connection with the objection to the specification, subject matter contained in the claims was “described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention,” as required under 35 U.S.C. § 112, first paragraph. In this regard, “subject matter of the claim need not be described literally ... in order for the disclosure to satisfy the description requirement.” Further, “The test for sufficiency of support in a parent application is whether the disclosure of the application relied upon ‘**reasonably conveys** to the artisan that the inventor had possession at that time of the later claimed subject matter.’ *Ralston Purina Co. v. Far-Mar-Co., Inc.*, 772 F.2d 1570, 1575, 227 USPQ 177, 179 (Fed. Cir. 1985) (quoting *In re Kaslow*, 707 F.2d 1366, 1375, 217 USPQ 1089, 1096 (Fed. Cir. 1983)).” (MPEP 2163.02.)

Further, the original claims of the specification as filed explicitly disclosed and claimed computer readable medium and computer program product in at least claims 11-15 and 24-27. Since the original claims are part of the specification and included at least “computer readable medium” and “computer program product,” the inventors at the time clearly had possession of subject matter contained in the claims. Accordingly, reconsideration and withdrawal of the rejections under 35 U.S.C. § 112, first paragraph are respectfully requested.

Claim Rejections Under 35 U.S.C. § 103: Claims 1-15 and 17-27

*I. Bases for Rejections are Unclear in the Office Action*

As a preliminary matter, the bases for rejections of claims 1-15 and claims 17-27 are unclear and the Office Action is internally inconsistent. In the “Response to Arguments” (p. 2), the Office Action states that the Assignee’s arguments for claims 1-15 and claims 17-24 were “**fully considered** and are **persuasive**.” However, the Office Action repeats nearly all of the rejections from the prior Office Action, for which the Assignee submitted arguments that were acknowledged as being “fully considered” by the PTO and found “persuasive.” In particular, the Office Action

repeats most of the rejections for claims 1-15 and the **same** rejections for claims 17-27 (p. 3). Therefore, the record is unclear as to whether and what arguments presented by Assignee were “fully considered” and found “persuasive.”

Further, although the rejections of claims 1-15 include a new reference (Shao), the Office Action is entirely unclear as to what reference is used as a basis for the rejection under § 103(a) (See MPEP 2143 B), or whether any reference is used as a basis (See MPEP 2143 A). The Office Action in the “Response to Arguments” indicates that newly cited reference Shao is used as the primary reference, and that the “teachings of Shao allow for a model creation system and can easily incorporate the techniques separating female from male data taught by Neti, Naito, and Kanevsky.” (p. 2) However, the Office Action also indicates Neti is used as the primary reference (p. 6). Therefore, it is entirely unclear to the Assignee what particular combination of art is being asserted against the claims (*i.e.*, what and whether a primary reference is being used and if so, how the primary reference is modified), and what teachings are believed to meet each of the limitations in the claims.

As discussed below, the Office Action, even though using multiple references, has not clearly identified where each and every element as set forth in the independent claims can be found in the cited references. Since all limitations, as set forth in the claims, have not been identified in the cited references, the Office Action has not even resolved the *Graham* factual inquiries. Therefore, the Office Action has not, at a minimum, produced a *prima facie* case of obviousness under § 103(a). Further, the Office Action has not provided a finding that each element, when used in an as yet undescribed combination, would “perform the same function as it does separately.” (See MPEP 2143 A)

In view of the foregoing, the Office Action fails to comply with guidelines set forth in the MPEP, and the rejections under § 103(a) should be withdrawn or revised to clearly address and resolve issues relating to Neti, Naito, Kanevsky, Shao, and Wark. At a minimum, the guidelines require, “In order to provide a complete application file history and to enhance the **clarity** of the prosecution history record, an examiner **must** provide clear explanations of all actions taken by the examiner during prosecution of an application,” and “If applicant’s arguments are persuasive, ... the examiner **must** provide in the next Office communication the reasons why the previous rejection is

withdrawn by referring specifically to the page(s) and line(s) of applicant's remarks which form the basis for withdrawing the rejection." (See MPEP, 707.07(f))

For at least the above reasons, the rejections under § 103(a) of claims 1-15 and 17-27 should be withdrawn and Assignee's arguments that were found persuasive clearly identified so that prosecution of the application can be advanced expeditiously.

#### Claim Rejections Under 35 U.S.C. § 103: Claims 1-15

Beginning at page 6, the Office Action rejects claims 1-16 under 35 U.S.C. § 103(a) as purportedly being unpatentable over Neti (U.S. Patent No. 5,953,701) in view of Naito (U.S. Patent No. 5,983,178) in further view of Kanevsky (U.S. Patent No. 6,529,902). Claim 16 was cancelled in a prior response and is no longer pending for examination. The Assignee respectfully traverses the rejections of claims 1-15, of which claims 1, 6, and 11 are independent.

#### *II. Claim 1*

A. None of the references disclose "determining a difference between each female phoneme model and each corresponding male phoneme model."

Claim 1 is directed to at least one computer readable medium encoded with instructions that, when executed by at least one processor, perform a method for generating a speech recognition model. Claim 1 recites, *inter alia* and in combination with other limitations of the claim, "determining a difference between each female phoneme model and each corresponding male phoneme model." Neti, which is relied on for disclosing this limitation, fails to disclose at least this limitation of claim 1.

The Office Action cites the Abstract, col. 3, ll. 37-49, and col. 4 ll. 10-29 as purportedly disclosing the above-cited limitation of claim 1. The cited sections have been reviewed, and this limitation as set forth in claim 1 is not found in the cited sections or elsewhere in Neti. The cited sections describe use of gender-dependent and context-dependent decision trees for determining a "level of gender dependence." (col. 4, ll. 53-54) There is, however, no disclosure anywhere in Neti of "determining a difference between each female phoneme model and each corresponding male phoneme model" as set forth in claim 1. The arguments in connection with this same issue as set forth by the Assignee in the responses filed November 23, 2009 and September 9, 2010 (which the

PTO found persuasive in both cases) still apply. Therefore, Neti fails to disclose at least the above-cited limitation for which Neti is relied upon.

None of Naito, Kanevsky, nor Shao cure the deficiencies of Neti. Naito is directed to clustering speakers based on vocal-tract configurations, and does not describe determining differences between female and corresponding male phoneme models. Kanevsky is directed to off-line detection of textual topical changes, and does not describe determining differences between female and corresponding male phoneme models. Shao is directed to apparatus and methods for “enabling non-expert users to generate, train, and test speech recognition models” (Shao, paragraph [0007]), and does not describe determining differences between female and corresponding male phoneme models. Accordingly, Neti, Naito, Kanevsky, and Shao considered alone or in combination fail to disclose at least “determining a difference between each female phoneme model and each corresponding male phoneme model,” as set forth in claim 1. For at least this reason, the rejection of claim 1 under 35 U.S.C. § 103(a) should be withdrawn.

B. None of the references disclose “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.”

The Office Action (p. 7) acknowledges that Neti fails to disclose “a gender-independent/dependent phoneme model” and purports that Naito discloses this aspect of the above limitation. However, as previously pointed out by the Assignee, Naito fails to disclose at least this limitation as set forth in claim 1. As noted above, Naito is directed to clustering speakers based on vocal-tract configurations. The Office Action mischaracterizes Naito at col. 15, line 54 – col. 16, line 25 in stating that it “teaches the recognition of phoneme dependent data which verifies whether data is independent of dependent, for example whether incoming data is within a range of a model or not.” (*sic*) There is nothing in this section about dependent or independent phonemes. There is also nothing in the section about “whether incoming data is within a range of a model or not.” The immediately preceding section (col. 15, lines 45-53) teaches that a “cluster model to be used in the phoneme verifier” has already been designated. Accordingly, there is nothing in Naito about “gender independent/dependent phoneme models” as asserted by the Office Action, for which Naito is relied upon. Further there is nothing in Naito about “creating a gender-independent phoneme

model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.”

Neither Kanevsky nor Shao cure the deficiencies of Naito. Kanevsky is directed to off-line detection of textual topical changes, but does not describe gender independent/dependent phoneme models. Shao is directed to apparatus and methods for enabling non-expert users to generate, train, and test speech recognition models, but does not describe gender independent/dependent phoneme models. Accordingly, Neti, Naito, Kanevsky, and Shao considered alone or in combination fail to disclose at least “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.” For at least this additional reason, the rejection of claim 1 under 35 U.S.C. § 103(a) should be withdrawn.

C. Shao fails to disclose “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value.”

The Office Action (p. 11) acknowledges that Neti, Naito, and Kanevsky fail “to teach creating a gender-independent phoneme model.” The Office Action cites FIG. 2, and paragraphs [0136] and [0137] of Shao as purportedly disclosing creating a gender-independent phoneme model. Paragraphs [0136] of Shao describes generally that different types of speech models could be created using the apparatus of Shao including “speaker independent or speaker dependent models.” Paragraph [0137] describes testing of speech models. Although the cited paragraphs broadly describe speech model creation, there is nothing about creating gender-independent phoneme models, let alone “creating a gender-independent phoneme model when the difference between the compared female phoneme model and the corresponding male phoneme model is less than a predetermined value,” as set forth in claim 1. Therefore, Shao fails to disclose the limitation for which Shao was relied upon. For this additional reason, the rejection of claim 1 under 35 U.S.C. § 103(a) should be withdrawn.

#### D. Conclusion regarding claim 1

At least in view of the foregoing reasons, Neti, Naito, Kanevsky, and Shao, considered alone or in combination, fail to disclose or suggest each and every limitation of claim 1. Therefore,

claim 1 patentably distinguishes over these cited references. Accordingly, withdrawal of the rejections of claim 1 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 2-5 depend from claim 1, and are therefore allowable for at least the same reasons.

### III. Claim 6

Claim 6 is directed to a system for generating speech recognition models and recites, *inter alia*, “determine a difference between each first phoneme model and each corresponding second phoneme model; create an independent phoneme model when the difference between the compared each first phoneme model and each corresponding second phoneme model is less than a predetermined value; and add, based upon a criterion, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.”

For reasons that should be clear from the discussion of claim 1 above in connection with Neti, Naito, Kanevsky, and Shao, these cited references fail to teach or suggest every element of claim 6. Therefore, claim 6 patentably distinguishes over Neti, Naito, Kanevsky, and Shao. Accordingly, withdrawal of the rejection of claim 6 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 7-10 depend from claim 6 and are therefore allowable for at least the same reasons.

### IV. Claim 11

Claim 11 encompasses a computer program product and recites, *inter alia*, “determine a difference between each first phoneme model and each second phoneme model; create an independent phoneme model when the difference between the each first phoneme model and the each corresponding second phoneme model is less than a predetermined value; and add, based on a criterion, one of the independent phoneme model, or both the first phoneme model and the corresponding second phoneme model to the speech recognition model.”

For reasons that should be clear from the discussion of claim 1 above in connection with Neti, Naito, Kanevsky, and Shao, these cited references fail to teach or suggest every element of claim 11. Therefore, claim 11 patentably distinguishes over Neti, Naito, Kanevsky, and Shao.

Accordingly, withdrawal of the rejection of claim 11 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 12-15 depend from claim 11 and are therefore allowable for at least the same reasons.

Claim Rejections Under 35 U.S.C. § 103: Claims 17-27

Beginning at page 24, the Office Action rejects claims 17-27 under 35 U.S.C. §103(a) as purportedly being unpatentable over Neti in view of Wark (U.S. Patent Publication No. 2003/0231775) and in further view of Naito. The cited references and the “Response to Arguments” have been reviewed, and the Assignee respectfully traverses the rejections of claims 17-27.

*V. Claim 17*

Claim 17 is directed to a computer readable medium product encoded with instructions and recites, *inter alia*, “computing best estimates that the current feature vector belongs to each one of the plurality of data classes; computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models.” At least these limitations of claim 17 are not taught or suggested in Neti, Wark, or Naito, and therefore claim 17 patentably distinguishes over Neti, Wark, and Naito. Reasons supporting this position are set forth below.

A. Neti fails to disclose “computing best estimates that the current feature vector belongs to each one of the plurality of data classes.”

The Office Action (p. 24) cites Neti at col. 3, ll. 50-67 and col. 5, ll. 9-21 as purportedly teaching “a first computing module configured to compute a current best estimates that the current feature vector belongs to one of the plurality of data classes.” (*sic*) This, however, is not what the claim recites. Claim 17 recites, “computing best estimates that the current feature vector belongs to



**each** one of the plurality of data classes.” (emphasis added) Col. 3, ll. 50-67 of Neti is directed to constructing context-dependent sub-phonetic models using binary questions in a decision-tree process. There is nothing in this section about computing best estimates that a current feature vector belongs to each one of a plurality of data classes. Col. 5, ll. 9-21 discloses that some analyzed phonemes exhibit gender dependence as represented in Fig. 3 of Neti. Fig. 3 shows the percentages each of analyzed consonants, semivowels, vowels, and diphthongs that exhibit gender dependence. Again, there is nothing in this section about computing best estimates that a current feature vector belongs to each one of a plurality of data classes. Therefore, contrary to the assertion in the Office Action, Neti fails to disclose “computing best estimates that the current feature vector belongs to each one of the plurality of data classes” for which Neti was relied upon. Therefore, the Office Action fails to identify where at least this limitation can be found in the cited references. For at least this reason, the rejection of claim 17 under 35 U.S.C. § 103(a) should be withdrawn.

**B. The limitations relating to “computing accumulated confidence values”  
are not identified by the Office Action nor disclosed by Neti, Wark, or Naito.**

The “Response to Arguments” broadly states that “Wark teaches the mathematical operations directed to computing accumulated confidence scores, whereby Wark improves the common and uncommon data sets of Neti.” (p. 3) The Office Action (p. 25) acknowledges that Neti fails to teach “computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models.” The Office Action (p. 25 – p. 27) then generally alludes to Wark by quoting several sections of Wark, and asserts “it would have been obvious ... to modify the system of Neti ... to incorporate” the above-identified limitations of claim 17 “as taught by Wark.”

In addition to the Office Action’s statement regarding Neti’s “data sets” being unclear and the assertion of obviousness being impermissibly conclusory (see MPEP 2142), the Office Action does not even attempt to point out where each of the above-cited limitations can be found in Wark. The sections of Wark ([0094], Fig. 4, [0129], [0130], and [0146]) generally alluded to in the Office

Action disclose a process for analyzing audio segments by parsing the segment into audio frames and audio clips and using extracted clip feature vectors in order to classify or identify an audio segment. However, there is no disclosure in the cited sections or elsewhere in Wark of “computing accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class; weighing the class-dependent phoneme models based on the accumulated confidence values; and recognizing the current feature vector based on the weighted class-dependent phoneme models,” as set forth in claim 17.

In particular, nowhere does Wark describe or suggest computing accumulated confidence values wherein “the confidence value for each data class based on the current best estimate and on previous confidence values for the each data class.” The Office Action has not even attempted to identify where this limitation can be found in Wark. Further, nowhere does Wark describe or suggest “weighing the class-dependent phoneme models based on the accumulated confidence values” or “recognizing the current feature vector based on the weighted class-dependent phoneme models,” as set forth in claim 17. The Office Action has not even attempted to identify where these limitations can be found in Wark. Therefore, Wark is lacking at least several limitations of claim 17 for which Wark was relied upon. For these additional reasons, the rejection of claim 17 under 35 U.S.C. § 103(a) should be withdrawn.

C. There is no relevance of Naito in the rejection.

The Office Action (p. 27) acknowledges that Neti and Wark fail to teach “creating a class-independent/dependent phoneme model,” and repeats text from the rejection of claim 1. However, claim 17 does not recite “creating a class-independent/dependent phoneme model.” Claim 17 is directed to recognizing speech, not creating class-independent/dependent phoneme models. Claim 17 does not recite “creating a class-independent/dependent phoneme model.” Accordingly, there is no relevance of Naito to the rejection of claim 17, and the rejection is in error and should be withdrawn for this additional reason.

Further, the Office Action seems to assert that Naito teaches "creating a class-independent/dependent phoneme model." However, the Office Action (p. 11) acknowledged in connection with claim 1 that Naito fails to teach this limitation. Accordingly, the Office Action is inconsistent at least in this respect.

D. Conclusion regarding claim 17.

At least in view of the foregoing, claim 17 patentably distinguishes over Neti, Wark and Naito, since these references (even if all relevant to the rejection, which Assignee does not concede) fail to disclose or suggest each and every limitation of the claim. Accordingly, withdrawal of the rejection of claim 17 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 18-20 depend from claim 17, and are therefore allowable for at least the same reasons.

*VI. Claim 21*

Claim 21 encompasses a system for recognizing speech data and recites, *inter alia*, "a first computing module configured to compute current best estimates that the current feature vector belongs to each one of the plurality of data classes; a second computing module configured to compute accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream; a weighing module configured to weigh the class-dependent phoneme models based on the accumulated confidence values; and a recognizing module configured to recognize the current feature vector based on the weighted class-dependent phoneme models."

For reasons that should be clear from the discussion of claim 17 above in connection with Neti, Wark, and Naito, these cited references also fail to teach or suggest every limitation of claim 21. Therefore, claim 21 patentably distinguishes over Neti, Wark, and Naito. Withdrawal of the rejection of claim 21 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 22-23 depend from claim 21 and are therefore allowable for at least the same reasons.

*VII. Claim 24*

Claim 24 encompasses a computer program product and recites, *inter alia*, “compute best estimates that the current feature vector belongs to each one of the plurality of data classes; compute accumulated confidence values for each of the plurality of data classes that the current feature vector belongs to each one of the plurality of data classes, the confidence value for each data class of the plurality of data classes based on the current best estimate for the data class and on previous confidence values for the data class, the previous confidence values associated with previous feature vectors of the audio stream; weigh the class-dependent phoneme models based on the accumulated confidence values; and recognize the current feature vector based on the weighted class-dependent ‘ phoneme models.”

For reasons that should be clear from the discussion of claim 17 above in connection with Neti, Wark, and Naito, these cited references also fail to teach or suggest every limitation of claim 24. Therefore, claim 24 patentably distinguishes over Neti, Wark, and Naito. Withdrawal of the rejection of claim 24 under 35 U.S.C. § 103(a) is respectfully requested.

Claims 25-27 depend from claim 24 and are therefore allowable for at least the same reasons.

General Comments on Dependent Claims

Since each of the dependent claims depends from a base claim that is believed to be in condition for allowance, for the sake of brevity, the Assignee believes that it is unnecessary at this time to argue the further distinguishing features of the dependent claims. However, the Assignee

does not necessarily concur with the interpretation of the previously presented dependent claims as set forth in the Office Action, nor does the Assignee concur that the basis for rejection of any of the previously presented dependent claims is proper. Therefore, the Assignee reserves the right to specifically address the further patentability of the dependent claims in the future.

**CONCLUSION**

A Notice of Allowance is respectfully requested. The Examiner is requested to call the undersigned at the telephone number listed below if this communication does not place the case in condition for allowance.

If this response is not considered timely filed and if a request for an extension of time is otherwise absent, Assignee hereby requests any necessary extension of time. If there is a fee occasioned by this response, including an extension fee, the Director is hereby authorized to charge any deficiency or credit any overpayment in the fees filed, asserted to be filed or which should have been filed herewith to our Deposit Account No. 23/2825, under Docket No. N0484.70762US00.

Dated: January 25, 2011

Respectfully submitted,

By

  
Marc S. Johannes

Registration No.: 64,978

WOLF, GREENFIELD & SACKS, P.C.

600 Atlantic Avenue

Boston, Massachusetts 02210-2206

617.646.8000